

UNIVERZA V LJUBLJANI  
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Nejc Rebernik

**Razvoj sistema za avtomatsko  
trgovanje na borzi**

DIPLOMSKO DELO

VISOKOŠOLSKI STROKOVNI ŠTUDIJSKI PROGRAM  
PRVE STOPNJE  
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: dr. Aleksander Sadikov

Ljubljana, 2017



COPYRIGHT. Rezultati diplomske naloge so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavo in koriščenje rezultatov diplomske naloge je potrebno pisno privoljenje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

*Besedilo je oblikovano z urejevalnikom besedil  $\text{\LaTeX}$ .*



Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Tematika naloge:

Kandidat naj preišče trenutne dobre prakse v tehničnem trgovanju in sestavi model za avtomatsko trgovanje. Cilj je izboljšati rezultat v primerjavi s strategijo nakupa in držanja. Za napovedovanje naj uporabi tehnične indikatorje. Kandidat naj tudi preizkusi možnost uporabe semantične analize besedila.



*Zahvaljujem se prof. dr. Aleksandru Sadikovu za strokovno pomoč in nasvete pri izdelavi diplomske naloge.*

*Posebej se zahvaljujem svojim domačim in prijateljem za podporo in spodbudo v času študija.*





# Kazalo

Povzetek

Abstract

<b>1</b>	<b>Uvod</b>	<b>1</b>
<b>2</b>	<b>Finančni trgi</b>	<b>3</b>
<b>3</b>	<b>Uporabljena orodja in tehnologije</b>	<b>5</b>
3.1	Razvojno okolje . . . . .	5
3.2	Zgodovinski podatki o cenah . . . . .	6
<b>4</b>	<b>Tehnični indikatorji</b>	<b>9</b>
4.1	Trend . . . . .	10
4.2	Zagon . . . . .	12
4.3	Volatilnost . . . . .	14
<b>5</b>	<b>Semantična analiza</b>	<b>17</b>
5.1	Vir podatkov . . . . .	17
5.2	Strokovna osnova . . . . .	18
<b>6</b>	<b>Avtomatsko trgovanje</b>	<b>19</b>
6.1	Filtri . . . . .	19
6.2	Prožilci . . . . .	22
6.3	Trgovalna strategija . . . . .	22

<b>7</b>	<b>Ovrednotenje</b>	<b>25</b>
7.1	Tehnično trgovanje . . . . .	25
7.2	Semantična analiza . . . . .	27
<b>8</b>	<b>Sklepne ugotovitve</b>	<b>29</b>
	<b>Literatura</b>	<b>31</b>

# Slike

3.1	JetBrains PyCharm integrirano razvojno okolje . . . . .	6
3.2	Podatki za AAPL med 1995 in 2005 . . . . .	8
4.1	Formula za izračun enostavnega gibajočega povprečja . . . . .	10
4.2	Gibajoče povprečje za 7 in 25 dni . . . . .	11
4.3	Izračun RSI . . . . .	12
4.4	Izračun RS . . . . .	12
4.5	Graf RSI . . . . .	13
4.6	Izračun stohastičnega oscilatorja . . . . .	14
4.7	Graf stohastičnega oscilatorja . . . . .	14
4.8	Formula za izračun standardne deviacije . . . . .	15
6.1	Trakovi gibajočega povprečja . . . . .	21
7.1	Primer grafa nakupov in prodaj . . . . .	26
7.2	Graf semantične ocene . . . . .	28



# Tabele

7.1	Primerjava trgovanja z referenčno oceno . . . . .	25
-----	---	----



# Seznam uporabljenih kratic

kratica	angleško	slovensko
<b>CSV</b>	comma separated values	vrednosti ločene z vejico
<b>RSI</b>	relative strength index	indeks relativne moči
<b>NYSE</b>	New York Stock Exchange	Newyorška borza
<b>USD</b>	United States dollar	Ameriški dolar
<b>SEC</b>	Securities and Exchange Com- mission	Komisija za vrednostne papirje in borzo





# Povzetek

**Naslov:** Razvoj sistema za avtomatsko trgovanje na borzi

**Avtor:** Nejc Rebernik

Diploma raziskuje načine napovedovanja gibanja cen vrednostnih papirjev na finančnih trgih in metode, ki brez zunanjega človeškega vnosa upravljajo s temi vrednostnimi papirji. Prvi del se ukvarja s kazalniki in vzorci s katerimi lahko napovedujemo gibanje cen. Drugi del se ukvarja s sprejemanjem trgovalnih odločitev glede na podane kazalnike. Skupaj tvorita avtonomen sistem za trgovanje na borzi.

**Ključne besede:** borza, vrednostni papirji, napovedovanje cen, avtomatsko trgovanje, tehnična analiza, semantična analiza besedila.



# Abstract

**Title:** Development of an automatic trading system

**Author:** Nejc Rebernik

The paper explores ways of predicting prices of securities in financial markets and methods that can autonomously manage these securities. The first part of this paper focuses on signals and indicators which we can use to predict price movements. The second part focuses on making trading decisions based on available indicators. Together they form a fully autonomous system for stock market trading.

**Keywords:** stock market, securities, price prediction, automated trading, technical analysis, semantic text analysis.



# Poglavje 1

## Uvod

Trgovanje z vrednostnimi papirji brez trgovalne strategije je podobno igranju rulete. S pomočjo računske moči, ki je danes na voljo v potrošniški strojni opremi lahko hitro analiziramo velike količine podatkov. Naš cilj je uporabiti metode tehnične in semantične analize za odkrivanje zakonitosti v podatkih in izdelati model za avtomatsko trgovanje. Glavna izziva s katerima se bomo soočili, sta napovedovanje cen in izdelava modela trgovalne strategije.

Za napovedovanje cen se bomo poslužili metode tehnične analize. Tehnična analiza temelji na zgodovinskih podatkih iz katerih izračunamo kazalnike, s katerimi lahko zgradimo trgovalni model.

Poleg tehnične analize bomo raziskali področje semantične analize. Za semantično analizo bomo uporabili podatke letnih finančnih poročil ter iz sprememb v jeziku in sentimentu poskušali odkriti povezavo z spremembami v ceni. Cilj je odkriti vzorce v besedilu, ki nakazujejo na rast ali padec cen vrednostnih papirjev v prihodnosti in uporabiti semantično analizo kot dopolnitev trgovalnega modela, ki temelji na tehnični analizi.

Največja motivacija je odkrivanje kazalnikov, ki nakazujejo na velike padce v cenah. Želimo odkriti vzorce, ki so predhodniki finančnih kriz in korekcij cen. S pravočasnim izhodom lahko preprečimo izgube, ki bi jih sicer utrpeli.

Dolgoročni trend trga je vedno pozitiven. Osredotočamo se na krajša,

deset do dvajsetletna obdobja v katerih se dogajajo največja nihanja. Dva primera velikih negativnih premikov v moderni zgodovini sta velika finančna kriza leta 1929 in velika finančna kriza leta 2008. S tehnično analizo zgodovinskih podatkov želimo odkriti vzorce, ki nakazujejo na naslednjo finančno krizo in se ji izogniti.

## Poglavje 2

# Finančni trgi

Izraz finančni trg je širok izraz za kakršenkoli trg, kjer se prosto izmenjujejo vrednostni papirji, dobrine in storitve. Glavna značilnost finančnih trgov je centralizirana lokacija za trgovanje brez ali z majhnimi provizijami. Cene izdelkov tako natančno odražajo tržno vrednost glede na ponudbo in povpraševanje.

Borze so tip finančnega trga, kjer se trguje z vrednostni papirji. Vrednostni papirji so lahko delnice, obveznice, investicijski kuponi vzajemnih skladov, indeksi vzajemnih skladov, naložbeni certifikati, nakupne opcije in drugi finančni produkti. V okviru naloge se bomo osredotočili samo na trgovanje z delnicami.

Pri kupovanju delnic se pojavi vprašanje katere delnice kupiti. Da omejimo problem bomo uporabljali podatke delnic, ki sestavljajo S&P 500 indeks. Indeks sestavlja 500 delnic iz različnih gospodarskih sektorjev podjetij, ki trgujejo na NYSE ali NASDAQ. Je eden izmen najbolj sledenih indeksov in se smatra za kazalnik moči ameriške borze in ekonomije.

Newyorška borza je bila ustanovljena in začela trgovati s prvimi vrednostnimi papirji leta 1792. Predstavlja največjo borzo v ZDA po volumnu trgovanja in drugo po številu vpisanih podjetij.

Glavna razlika med Newyorško borzo in NASDAQ je v lokaciji trgovanja. Na Newyorški borzi se vse transakcije zgodijo preko ljudi na borznem

parketu, medtem ko se na NASDAQ vse transakcije v celoti opravijo elektronsko. Newyorška borza je tipično asociirana z starejšimi, bolj uveljavljenimi podjetji, medtem ko je NASDAQ tipično povezan z novejšimi tehnološkimi podjetji z agresivno rastjo in višjo volatilnostjo. Kljub razlikam, borza na kateri kotira delnica ni parameter, ki se ga upošteva pri trgovanju.

Trgovanje na obeh borzah poteka od ponedeljka do petka med 9.30 in 16.00 uro po vzhodnem standardnem času. Pred odprtjem in po zaprtju rednega prometa je v oknu 4 ur možno tudi elektronsko trgovanje. Trgovanje v teh urah se razlikuje od trgovanja v rednem času zaradi višje volatilnosti in manjše likvidnosti. Trgovanje je prekinjeno tudi med državnimi prazniki, kadar so spremenjeni tudi časi začetka in konca trgovanja.



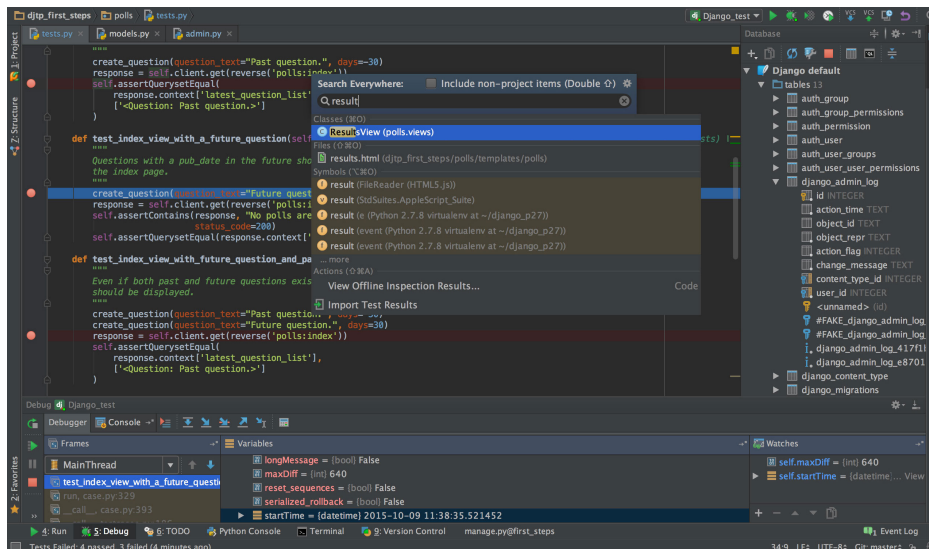
## Poglavje 3

# Uporabljena orodja in tehnologije

### 3.1 Razvojno okolje

Za analizo podatkov bomo uporabili programski jezik Python. Omogoča hiter razvoj in prototipiranje zaradi velikega števila knjižnic za delo s podatki. Za potrebe tega projekta bomo uporabili paket Anaconda, ki vsebuje distribucijo Python 3.5 in knjižnice za delo z velikimi podatki, analiziranjem in napovedovanjem. Vključuje več kot 100 paketov iz področja podatkovnega rudarjenja in matematike. Med bolj pomembnimi za nas so numpy za shranjevanje in računske operacije nad vektorji in tabelami ter pyplot za vizualizacijo podatkov in risanje grafov.

Za podporo pri razvijanju bomo uporabili integrirano razvojno okolje JetBrains PyCharm. Pomaga pri razvoju s predlogi dopolnjevanja kode, sintaktično analizo in načinom za razhroščevanje. Način za razhroščevanje omogoča premikanje po korakih v programski kodi. Tako hitreje odkrijemo napake v programu in pospešimo razvoj.



Slika 3.1: JetBrains PyCharm integrirano razvojno okolje

## 3.2 Zgodovinski podatki o cenah

Za zgodovinske podatke o delnicah bomo vzeli seznam podjetij iz indeksa S&P 500. V indeks je vključenih 500 najuspešnejših podjetij, ki kotirajo na NYSE ali NASDAQ. Vključuje podjetja iz različnih sektorjev in predstavlja splošno ekonomsko klimo ameriške ekonomije. Seznam podjetij na indeksu se redno spreminja. Seznam, ki ga uporabljamo prikazuje stanje na dan 29.6.2017. Za vsako podjetje na seznamu bomo iz Google Finance izvozili zgodovinske podatke za učno in testno množico. Učna množica obsega dni med 29.6.1995 in 29.6.2005, testna množica pa med 29.6.2006 in 29.6.2016. Za analizo bomo uporabili samo podjetja, ki kotirajo oziroma imajo dostopne podatke od 29.6.1995. Če ni na voljo podatkov za celotno časovno obdobje, podjetja ne bomo vključili v analizi.

Za vsako delnico obstaja enoličen identifikator. Sestavljen je iz kratice borze na kateri kotira in imena delnice, ki sta med seboj ločena z dvopičjem. Primer za Microsoft: NASDAQ:MSFT. Podatki, ki smo jih pridobili iz Google Finance vsebujejo ceno ob odprtju trga, dnevni minimum in maksimum,

cena ob zaprtju trga in količina trgovanja. Te parametre bomo uporabili za izračun kazalnikov in tehnično analizo.

Portal Google Finance omogoča izvoz zgodovinskih v obliki CSV. Ker bi bil ročni prenos vseh 500 datotek preveč zamuden smo napisali skripto za prenašanje podatkov. Skripta se poslužuje enotne oblike povezav. Vsebovati mora identifikator delnice, začetni datum, končni datum in oznako da želimo prenesti CSV datoteko.

```
https://www.google.com/finance/historical?q={}  
&startdate={}&enddate={}&output=csv
```

Na mestih {} vstavimo ustrezne podatke. Primer povezave za Apple Inc. od 29. junija 1995 do 29. junija 2005:

```
https://www.google.com/finance/historical?q=AAPL  
&startdate=29+Jun%2C+1995&enddate=29+Jun%2C+2005&output=csv
```

Podatki so shranjeni kot surovo besedilo. Vsaka vrstica vsebuje datum, ceno ob začetku trgovanja, dnevni maksimum, dnevni minimum, ceno ob koncu trgovanja in volumen. Stolpci so med seboj ločeni z vejicami. Podatki so urejeni po času padajoče. Primer prvih 10 vrstic za zgornji URL:

1	Date, Open, High, Low, Close, Volume
2	29-Jun-05, 5.32, 5.33, 5.16, 5.20, 113703681
3	28-Jun-05, 5.36, 5.37, 5.31, 5.33, 87908072
4	27-Jun-05, 5.26, 5.44, 5.24, 5.30, 150429797
5	24-Jun-05, 5.58, 5.59, 5.38, 5.39, 115544828
6	23-Jun-05, 5.55, 5.68, 5.52, 5.56, 168681695
7	22-Jun-05, 5.47, 5.51, 5.45, 5.51, 107428251
8	21-Jun-05, 5.39, 5.46, 5.34, 5.41, 92792819
9	20-Jun-05, 5.41, 5.44, 5.35, 5.37, 81327953
10	17-Jun-05, 5.50, 5.51, 5.40, 5.47, 150353784
11	16-Jun-05, 5.31, 5.44, 5.26, 5.43, 138585951

Slika 3.2: Podatki za AAPL med 1995 in 2005

## Poglavje 4

# Tehnični indikatorji

Za napovedovanje premikov cen v prihodnosti potrebujemo kazalnike, ki iz gibanja cen v preteklosti pridobijo dodatno informacijo. Obstajajo štiri glavne kategorije indikatorjev:

- trend
- zagon
- volatilnost
- količina

Za kvalitetno napovedovanje moramo uporabiti indikatorje iz večih kategorij. Kopičenje indikatorjev iz iste kategorije zmanjšuje verodostojnost ocene. Ker se indikatorji iz iste kategorije opirajo na iste podatke, kombiniranje ne prinese dodatne informacije.

Indikatorji ki so odvisni od časa so parametrizirani s časovnim oknom pogleda v preteklost. Odvisno od velikosti časovnega okna se spreminja odzivnost na premike cen. Indikatorji z dolgim časovnim oknom so primerni za zaznavanje dolgoročnih trendov, medtem ko so indikatorji s kratkim časovnim oknom primerni za odkrivanje kratkoročnih trendov.

## 4.1 Trend

Trend je generalna smer cene delnice za določeno obdobje. Trend je lahko pozitiven, negativen ali nevtralen. Za izračun trenda obstaja več metod.

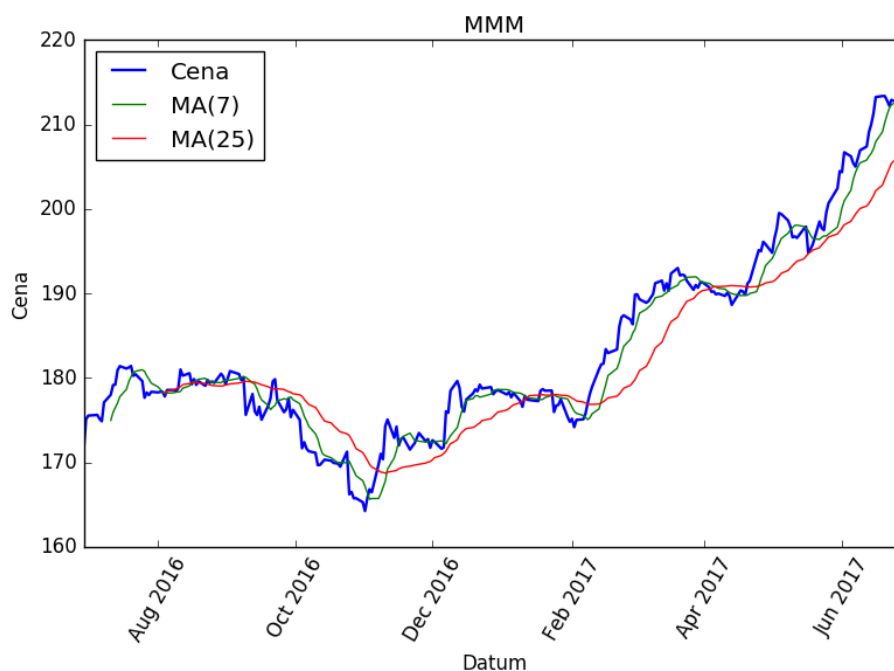
### Gibajoče povprečje

Enostavno gibajoče povprečje je najbolj pogosta oblika trenda. Parametrizirano je s številom dni vključenih v izračun.

$$MA(n) = \frac{\sum_{i=1}^n a_i}{n}$$

Slika 4.1: Formula za izračun enostavnega gibajočega povprečja

Gibajoče povprečje ima zamik za cenami. Ko cena delnice naraste, gibajoče povprečje z zamikom sledi rasti, ker v svojem izračunu še vedno upošteva stare vrednosti. Večji kot je časovni okvir gibajočega povprečja, večji je zamik trendne črte. Poglejmo primer gibanja cen za 365 dni, z gibajočim povprečjem za 7 in 25 dni.



Slika 4.2: Gibajoče povprečje za 7 in 25 dni

Iz slike lahko razberemo da gibajoče povprečje s 7-dnevnim oknom hitreje sledi spremembam v ceni in se bolj tesno prilega obliki grafa. Na drugi strani ima gibajoče povprečje s 25-dnevnim oknom opazno večji zamik in predstavlja bolj dolgoročne trende. Zamik trendne črte je v žargonu poznan kot „lag“. Z daljšim gibajočim povprečjem bolj zaostajamo za spremembami in ne zaznavamo majhnih premikov. Počasna odzivnost gibajočega povprečja ni nujno negativna. Počasna trendna črta nam lahko pomaga zgladiti majhne korekcije cene v trgu. Če trgujemo v velikostnem redu tednov, potem si želimo trendno črto s hitro odzivnostjo. Nasprotno, če trgujemo v velikostnem redu let, potem si želimo počasnejšo trendno črto z večjim časovnim okvirjem. Časovni okvir gibajočega povprečja moramo prilagoditi trgovalni strategiji.

## 4.2 Zagon

Zagon je merilo hitrosti s katero se spreminja cena delnice. Večji zagon pomeni hitrejša spremembe v ceni. Uporaba zagona je koristna za odkrivanje precejenjenih in podcejenjenih premikov.

### Indeks relativne moči

Indeks relativne moči je zagonski indikator, ki je uporabljen za ocenjevanje precejenjenosti in podcejenjenosti delnic. V podanem časovnem okviru primerja vsa višanja in nižanja cene. Tipično časovno okno za indeks relativne moči je 14 dni.

$$RSI(n) = 100 - \frac{100}{1 + RS}$$

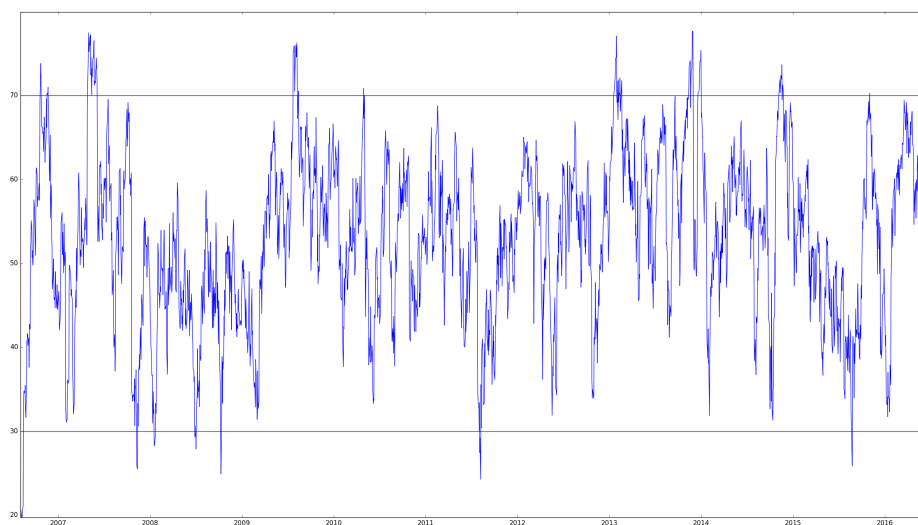
Slika 4.3: Izračun RSI

$$RS = \frac{\frac{AvgGain * (n-1) + Gain}{n}}{\frac{AvgLoss * (n-1) + Loss}{n}}$$

Slika 4.4: Izračun RS

V enačbi je  $n$  število dni v časovnem oknu. Indeks lahko računamo po  $n$  dneh od začetka naših podatkov. Po  $n$  dneh je povprečna pridobitev in povprečna izguba vsota pridobitev in vsota izgub  $n$  dni deljena z  $n$ . Za vsak naslednji dan povprečje izračunamo tako da vzamemo povprečje prejšnjega dne, ga pomnožimo z  $n-1$ , prištejemo pridobitev oziroma izgubo in delimo z  $n$ .





Slika 4.5: Graf RSI

Indeks relativne moči je omejen med 0 in 100. Tipična nastavitev nivojev za zaznavanje preценjenosti je 30 in 70. Če je indeks večji kot 70, je to znak preценjenosti delnice in je signal za prodajo. Če je indeks manjši kot 30, je to znak podценjenosti delnice in je signal za nakup. Mejo lahko prilagajamo za večjo ali manjšo občutljivost. Če višamo zgornji prag od 70 in navzgor, bomo kasneje zaznali preценjenost. Če nižamo spodnji prag od 30 navzdol, bomo kasneje zaznali podценjenost. Zgornjo in spodnjo mejo lahko spreminjamo neodvisno. Če želimo kasneje zaznati preценjenost potem dvignemo zgornjo mejo na 80, spodnjo mejo za zaznavanje podценjenosti pa pustimo na 30.

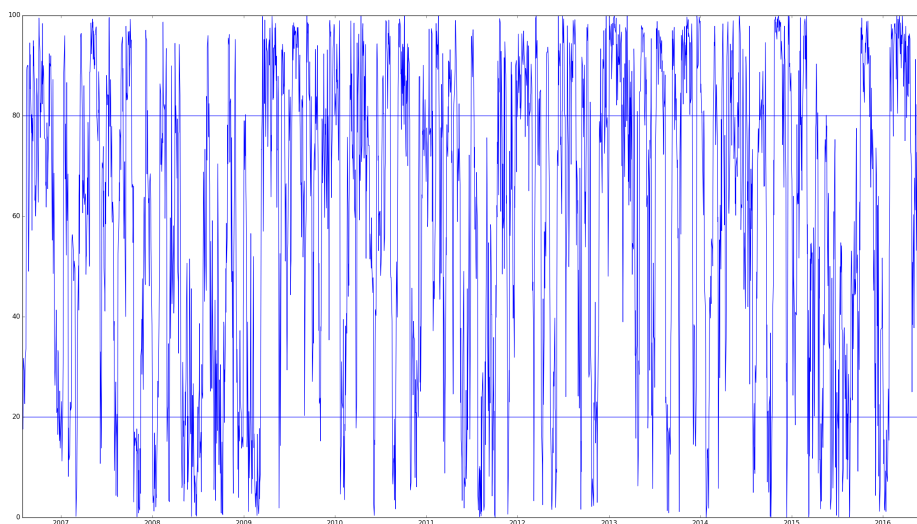
## Stohastični oscilator

Stohastični oscilator je zagonski indikator, ki je uporabljen za ocenjevanje preценjenosti in podценjenosti delnic. V podanem časovnem okviru primerja razliko med najnižjo in najvišjo ceno.

$$SO_n = \frac{(price - \min_n)}{(\max_n - \min_n)} * 100$$

Slika 4.6: Izračun stohastičnega oscilatorja

Stohastični oscilator izračunamo na podlagi dnevnih minimumov in maksimumov ter cene ob koncu trgovalnega dne za določeno časovno okno. Za izračunane vrednosti izračunamo še 3-dnevno gibajoče povprečje. Ko 3-dnevno gibajoče povprečje stohastičnega oscilatorja preseka čez vrednost stohastičnega oscilatorja je to signal za upadajoč trend in prodajo. Obratno, ko 3-dnevno povprečje preseka pod vrednost stohastičnega oscilatorja je to signal za naraščanje trenda in nakup.



Slika 4.7: Graf stohastičnega oscilatorja

### 4.3 Volatilnost

Volatilnost je statistična mera razpršenosti vrednosti. Visoka volatilnost pomeni večja nihanja cen. Nizka volatilnost pomeni manjša nihanja cen. Če je naša strategija naklonjena tveganju, potem je visoka volatilnost zaželena.

## Standardna deviacija

Standardna deviacija je ena izmed osnovnih statističnih mer, ki jo uporabimo tudi za izračun volatilnosti. Za izračun volatilnosti uporabimo standardno deviacijo.

$$\sigma = \frac{\sum_{i=1}^n (c_i - \bar{c})^2}{n}$$

Slika 4.8: Formula za izračun standardne deviacije

S spremenljivko  $n$  določamo časovno okno v katerem merimo standardno deviacijo.



# Poglavje 5

## Semantična analiza

### 5.1 Vir podatkov

Po zvezni zakonodaji morajo vsa javno trgovana podjetja v ZDA vsako leto izdati letno poročilo „Form 10-K“ s finančnim izkazom poslovanja, ki vsebuje pregled poslovanja, podatke o skupnem kapitalu s katerim razpolaga, izbrane finančne kazalnike, bonuse izplačane članom uprave ter druge pravne dokumente. Poleg finančnega poročila morajo podjetja izdati tudi letno poročilo delničarjem. Vsebuje nagovor generalnega izvršnega direktorja, finančne podatke, načrte novih produktov, delovanje hčerinskih podjetij ter aktivnosti na področju raziskav in razvoja.

Za semantično analizo se bomo osredotočili na letno poročilo delničarjem, še bolj natančno na nagovor generalnega izvršnega direktorja. Poleg napovedovanja na podlagi tehničnih indikatorjev se bomo lotili analize letnih poročil. Nad vsakim letnim poročilom bomo izvedli semantično analizo besedila in primerjali sentiment besedila z rastjo ali padcem delnice v naslednjem letu. Iskali bomo vzorce, pogostost besed, uporabljene besede in druge značilnosti, ki lahko kažejo na klimo v podjetju in moč razvoja ali stagniranja.

## 5.2 Strokovna osnova

Ideja sloni na osnovi nevro-lingvistične analize, ki pravi, da besede in način tvorjenja stavkov vsebujejo informacijo o resničnem prepričanju. Teorija trdi, da če se oseba zavestno laže ali poskuša prikrojiti podatke se to podzavestno manifestira v načinu izražanja. Naš cilj je iskanje povezave med sentimentom besedila in spremembo cen v prihodnosti. Predpostavljamo, da je povečan negativni sentiment povezan s slabšim poslovanjem v prihodnosti in povečan pozitiven sentiment z boljšim poslovanjem v prihodnosti. Iskali bomo semantične značilnosti in povezavo teh značilnosti s premiki cen.

Za semantično ocenjevanje smo naložili slovar pozitivnih in negativnih besed sestavljen za namene raziskave odkrivanja lažnih recenzij [8]. Vsebuje slovar 2005 pozitivnih in 4782 negativnih besed.

Za učenje modela smo naložili letna poročila dvajsetih podjetij iz S&P 500 od leta 1996 do 2016. Iz letnih poročil smo vzeli besedilo pisem delničarjem. Učno množico predstavljajo dokumenti od leta 1996 do 2009, testno množico pa dokumenti od leta 2010 do 2016. Za ocenjevanje smo uporabili slovar besed s pozitivnim in negativnim sentimentom. Učenje modela smo izvajali za vsako podjetje posebej, ker ima vsaka oseba svoj slog pisanja in bi mešanje različnih slogov pisanja zakrilo spremembe. Za vsako podjetje smo za vsako leto izračunali semantične kazalnike. Ti kazalniki so:

- odstotek semantično obarvanih besed
- odstotek pozitivno obarvanih semantičnih besed
- odstotek negativno obarvanih semantičnih besed
- razmerje med pozitivno in negativno semantično obarvanimi besedami

## Poglavje 6

# Avtomatsko trgovanje

Za popolnima avtomatiziran trgovalni sistem je potrebno imeti izdelano trgovalno strategijo. Trgovalna strategija je sestavljena iz 3 glavnih delov:

- filtri
- prožilci
- pravila

### 6.1 Filtri

Filtri so pogoji, ki jih določimo, da se začnemo zanimati za nakup delnice. Filter je lahko aktiven ali neaktiven. Filter ni signal za nakup, temveč samo predhoden pogoj signalu za nakup. Filter je lahko aktiven več dni zapored, dokler je pogoj za aktivnost filtra izpolnjen.

Filter je lahko sestavljen iz več različnih pogojev. Opazujemo lahko vrednost določenega kazalnika glede na podan prag. Lahko primerjamo dva enaka kazalnika z različnimi parametri, kot je na primer enostavno gibajoče povprečje z različnimi vrednostmi časovnega okna.

Filter ne more samostojno prožiti signala za nakup ali prodajo. Za proženje signala potrebujemo kombinacijo filtra in prožilca.

## **Enostavno gibajoče povprečje**

Z enostavnim gibajočim povprečjem lahko sestavimo več filtrov v kombinaciji s ceno, drugim enostavnim gibajočim povprečjem ali kateremkoli od drugih indikatorjev. Začnimo z najbolj enostavnim.

### **Enostavno gibajoče povprečje in cena**

Enostavno gibajoče povprečje s katerimkoli parametrom časovnega okna se v vsakem trenutku nahaja nad ali pod trenutno ceno delnice. Če se nahaja pod trenutno ceno, pomeni da cena raste hitreje glede na gibajoče povprečje obdobja in je „podprta“. Obratno, če se gibajoče povprečje nahaja nad ceno, pomeni da cena pada hitreje kot povprečje prejšnjega obdobja in ima „upor“ glede na gibajoče povprečje. To je najbolj enostaven in pogosto uporabljen filter.

### **Različno parametrizirani gibajoči povprečji**

Za dve gibajoči povprečji z različnim parametrom časovnega okna se v vsaki časovni enoti eno nahaja nad drugim. Filter tako ustvarimo na pogoju katero gibajoče povprečje je višje. Ko je krajše gibajoče povprečje nad daljšim, je to filter za nakup. Nasprotno, ko je daljše gibajoče povprečje nad krajšim, je to filter za prodajo.

### **Razmerje rasti gibajočih povprečij**

Odvod gibajočega povprečja nam vrne hitrost rasti oziroma padanja. Krajše gibajoče povprečje izgubi hitrost rasti preden pade pod počasnejše gibajoče povprečje. Točka, kjer hitrost rasti krajšega gibajočega povprečja pade pod hitrost rasti daljšega gibajočega povprečja nakazuje možnost padanja. Zaradi narave preseka dveh črt bi lahko razmerje rasti uvrstili med prožilce, vendar bi zaradi narave majhnih korekcij v cenah sprožilec deloval preveč sporadično. Namesto kot prožilec, je smiselna uporaba kot filter, ki je aktiven, ko krajše



gibajoče povprečje raste z manjšo hitrostjo kot daljše gibajoče povprečje v kombinaciji z enim od prožilcev.

### Trakovi gibajočih povprečij

Gibajoče povprečje je parametrizirano s številom dni za katere računamo povprečje. Če izračunamo gibajoča povprečja za različna obdobja dobimo trakove, ki se z različnimi hitrostmi premikajo za ceno.



Slika 6.1: Trakovi gibajočega povprečja

Iz grafa lahko vidimo, da so med padanjem cen trakovi večinoma nad ceno in označujejo upor. Dokler so vsi trakovi nad ceno lahko pričakujemo nadaljnje padanje cen. Ko začne cena rasti sčasoma vsi trakovi gibajočega povprečja pridejo pod ceno in podpirajo rast. Zelo površno lahko odstotek trakov, ki podpirajo ceno smatramo kot kazalnik verjetnosti obrata trenda. Če so vsi trakovi pod trenutno ceno je podpora za rast zelo močna. Obratno, če so vsi trakovi nad ceno je upor zelo močen in kaže da bo cena še naprej padala.

## 6.2 Prožilci

Prožilci so enkratni dogodki v času. Uporabljajo se za generiranje signalov nakupa in prodaje v sodelovanju s filtri. Nasprotno od filtra, prožilec nikoli ne traja več kot 1 časovno enoto. Prožilec je aktiviran, če je ob prehodu iz prve v drugo časovno enoto izpolnjen pogoj. Prožilec se poleg potrditve filtra in proženja nakupa lahko uporabi tudi za preprečevanje vstopa v slab nakup.

Za razliko od filtra lahko prožilec sproži signal za nakup ali prodajo brez podpore filtra. Takšne scenarije lahko določimo v svojih trgovalnih pravilih.

### Presek enostavnega gibajočega povprečja

Tako kot filter, je tudi najbolj enostaven prožilec sestavljen iz enostavnega gibajočega povprečja.

Dve gibajoči povprečji z različnim časovnim oknom lahko ustvarita prožilec ko se črti presekata. V trenutku ko hitrejša gibajoča povprečja naraste nad počasnejša gibajoča povprečja dobimo prožilec za nakup. Obratno, trenutek ko hitrejša gibajoča povprečja pade pod počasnejša gibajoča povprečja ustvari prožilec za prodajo.

### Precejenost in podcejenost

Precejenost in podcejenost se ocenjuje z indeksom relativne moči in stohastičnim oscilatorjem. V primeru prestopa praga precejenosti indikatorja ustvarita prožilec za prodajo, v primeru padca pod prag podcejenosti pa prožilec za nakup.

## 6.3 Trgovalna strategija

Trgovalna strategija je vse zajemajoč sistem za pravil in pogojev za vstop in izstop iz nakupa ter upravljanje denarja in tveganja. Pravila za vstop

in izstop iz nakupa se sestavlja iz kombinacije filtrov in prožilcev. Trgovna strategija ne narekuje samo nakupa in prodaje ampak tudi upravljanje denarja med trgovanjem. Več pogojev v trgovni strategiji pokrije več trgovnih scenarijev, ampak hkrati izgubi splošnost v novih scenarijih in se pretirano prilagodi enkratnim pojavom v zgodovini.

### 6.3.1 Vstop in izstop iz nakupa

Prva veja trgovne strategije - vstop in izstop - se ukvarja s pravili in scenariji za nakup in prodajo delnic. Kombinacija filtrov in prožilcev je strategija za nakup in prodajo. Poleg same kombinacije filtrov in prožilcev na rezultat pomembno vplivajo tudi parametri kazalnikov. Dva primera parametrov kazalnikov sta časovno okno za enostavno gibajoče povprečje in meja precejenosti in podcejenosti za indeks relativne moči. Poleg kombinacije filtrov in prožilcev na rezultat vplivajo tudi parametri kazalnikov.

Strategija za nakup je sestavljena iz trakov gibajočega povprečja in indeksa relativne moči. Uporabljamo šest trakov s parametri časovnega okna 50, 100, 150, 200, 250 in 300. Pri nakupu iščemo podporo cene čim večih trakov. Z indeksom relativne moči preverjamo ali je nakup v določenem trenutku smiseln. V fazi spremembe iz negativnega v pozitiven trend se pojavijo velika kratkoročna nihanja v ceni. S pogojem da je indeks relativne moči manjši od 50 se zavarujemo pred nakupom po kratkoročno precenjeni ceni.

Strategija za prodajo uporablja kriterij podpore 300-dnevnega gibajočega povprečja in odvod 50-dnevnega gibajočega povprečja in faktor rasti. Ko cena izgubi podporo 300-dnevnega gibajočega povprečja in je rast 50-dnevnega povprečja že negativna imamo predpogoj za prodajo. Da se izognemo nepotrebnim prodajam v času horizontalnega trenda uporabimo koeficient spremembe cene. Za obdobje 300 dni izračunamo razliko med ceno. Če je sprememba manjša od odstotka, ki ga določimo, potem prodaje ne izvedemo.

### 6.3.2 Upravljanje z denarjem

Druga veja trgovalne strategije se ukvarja z upravljanjem denarja. Določa razmerje delitve denarja za vlaganje v različne delnice. Zagotavlja, da smo vedno dovolj likvidni, da lahko kupimo nove delnice, če se pojavi ugoden pogoj za nakup. Z razpršenostjo delnic preprečuje, da v primeru prodaje pod nakupno ceno izgubimo čim manjši delež celotnega kapitala.

# Poglavje 7

## Ovrednotenje

### 7.1 Tehnično trgovanje

Za učenje in ovrednotenje modela tehničnega trgovanja smo uporabili testno in učno množico v razmerju 50:50. Obe vključujeta podatke za 10 let. Učna množica zavzema podatke od 29. junija 1995 do 29. junija 2005. Testna množica zavzema podatke od 29. junija 2006 do 29. junija 2016. Med njima je enoletna razlika, da preprečimo prekrivanje podatkov. Glavni cilj je bil zaznavanje in izogibanje velikim izgubam v času finančnih kriz. V učni množici je zajet tako imenovan "dot com" mehurček iz začetka dvatisočih, v testni množici pa velika recesija med letoma 2007 in 2009.

Tabela 7.1: Primerjava trgovanja z referenčno oceno

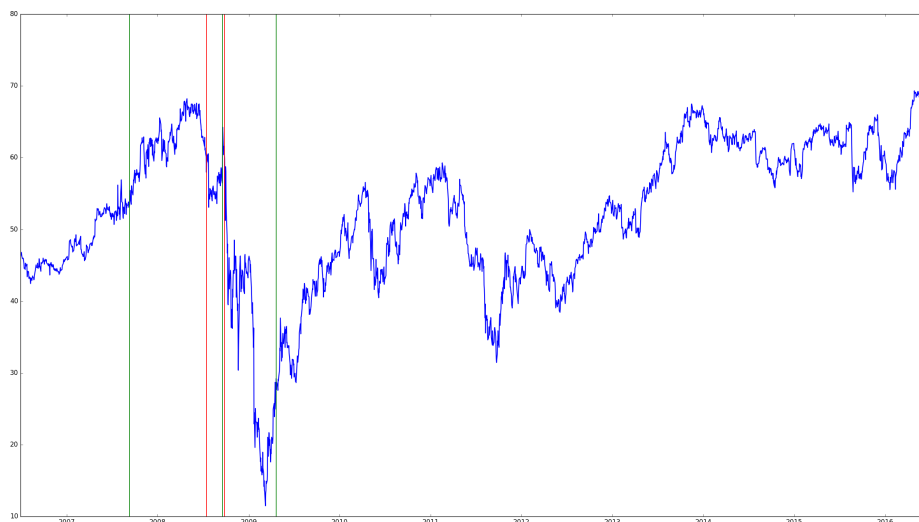
	Začetni vložek	Vsota	Dobiček
Referenčna ocena	250.000,00\$	467.670,71\$	217.670,71\$ (87,06%)
Trgovanje	250.000,00\$	563.783,74\$	313.783,74\$ (125,51%)

Za ovrednotenje rezultatov modela potrebujemo referenčno oceno. Primerna referenčna ocena je dobiček pri nakupu ob začetku in prodaja na koncu 10-letnega trgovalnega obdobja. Začetni vložek v višini 250.000,00 USD se pri nakupu 29. junija 2006 in prodaji 29. junija 2016 poveča za 217.670,71 USD (87,06%), za končni znesek 467.670,71 USD. Ta znesek je referenčna

ocena, ki jo mora naš model preseči, da se smatra kot uspešen. Večji kot je presežek referenčne ocene, boljši je model.

Z modelom smo presegli referenčno oceno. V istem obdobju smo začetni vložek 250.000,00 USD povečali za 313.783,74 USD(125,51%), za končni znesek 563.783,74 USD. Glede na referenčno oceno smo dosegli izboljšavo 20,55%.

Rezultat je uspešen predvsem zaradi uspešnih prodaj pred zaznanimi velikimi padci cen. Primer je viden na grafu nakupov in prodaj. Modra črta prikazuje ceno, zelena navpična črta prikazuje mesto nakupa, rdeča navpična črta pa mesto prodaje.



Slika 7.1: Primer grafa nakupov in prodaj

Prvotno je model pretirano reagiral na premike v obdobjih z majhnimi spremembami in zaradi tega izgubljal denar. Eden izmed ključev za uspeh je zanemarjanje majhnih sprememb. Pred prodajo pogledamo enoletno razliko v spremembi 300-dnevnega enostavnega gibajočega povprečja. Če se je sprememba navzgor ali navzdol manjša kot določeno število odstotkov, delnice ne prodamo. S tem rešimo težavo v času horizontalnega trenda in prodamo samo ob velikih spremembah.

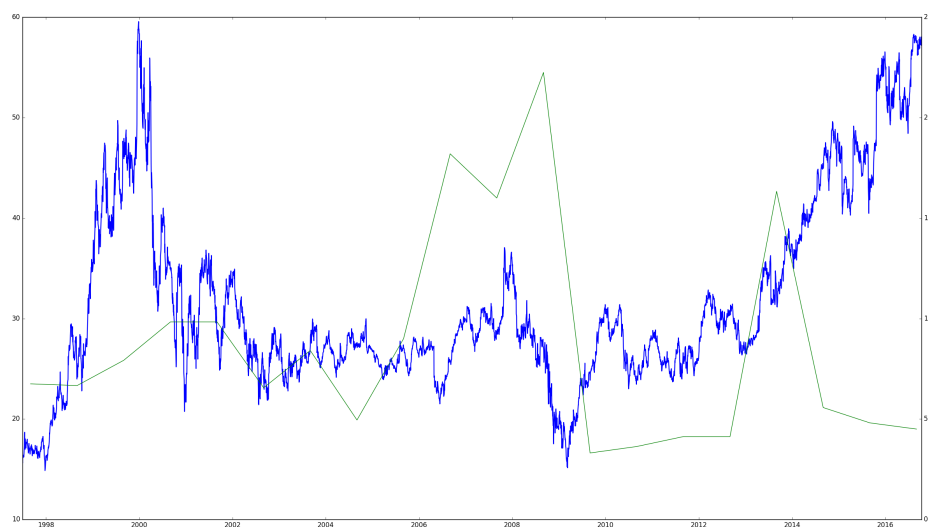
Odstotek spremembe predstavlja kompromis med izkoriščanjem prodajnih priložnosti in zanesljivostjo. Majhen prag odstotka spremembe je bolj nagnjen k prodaji ob manjših spremembah. Visok prag odstotka spremembe je bolj konservativen in potrebuje večjo spremembo v ceni za prodajo. Na razponu od 0 do 100, se po testiranju optimum giba med 30 in 60 odstotki. Vrednosti bližje 0 povzročajo težavo v času horizontalnega trenda, kjer kupujejo in prodajajo delnice z minimalno spremembo v ceni, pogosto tudi z izgubo. Vrednosti bližje 100 zmanjšujejo verjetnost za prodajo. V skrajnem primeru se izkaže kot nedejavnost, kjer se prodaja nikoli ne izvede, ker je sprememba v ceni premajhna.

## 7.2 Semantična analiza

Napovedovanje rezultatov semantične analize je bolj podobno tradicionalnemu strojnemu učenju. Za podatke o sentimentu smo napovedali ali bo cena delnice čez eno leto višja ali nižja. Referenčna ocena za testno množico je klasifikacijska točnost z večinskim klasifikatorjem.

Rezultati, ki smo jih dobili z semantično analizo so v najboljšem primeru enaki večinskemu klasifikatorju. Za izboljšanje rezultatov smo poskusili uporabiti večletni trend spremembe cen in sentimenta. Učna in testna množica sta se spremenili, a nismo dobili boljših rezultatov.

Za boljšo predstavbo smo za Microsoft(MSFT) vizualizirali gibanje cene delnice ter razmerje med pozitivno in negativno obarvanimi besedami v letnih poročilih. Modra črta predstavlja gibanje cene delnice med leti 1998 in 2016. Zelena črta predstavlja razmerje med pozitivnimi in negativnimi besedami v letnih poročilih. Vrednost deset pomeni, da je pozitivnih besed desetkrat toliko kot negativnih. Višja vrednost pomeni bolj pozitivno obarvano besedilo.



Slika 7.2: Graf semantične ocene

Če pogledamo graf gibanja sentimenta skozi čas lahko opazimo visok skok v pozitivno smer pred finančno krizo leta 2008 in strm padec po krizi. Padec sentimenta je v tem primeru zelo izrazit. Podoben vzorec se ponovi leta 2013 z visokim porastom v pozitivnem sentimentu in padcu naslednje leto, vendar brez spremembe v dolgoročnem trendu cene. Dejstvo da se negativen sentiment izkaže šele po spremembi cene je ključni problem v napovedovanju.



## Poglavje 8

### Sklepne ugotovitve

Na področju tehničnega trgovanja smo razvili model, ki za 20,55% izboljša dobiček pri trgovanju v primerjavi z referenčno oceno. Model uspešno zaznava dolgoročne spremembe trendov in vzorce pred velikimi padci cen. Z uspešnim zaznavanjem padcev in rasti model samodejno kupuje in prodaja delnice.

Semantična analiza z semantičnimi kazalniki ni pokazala uporabnih rezultatov. Na testni množici smo dobili rezultate, ki so enaki ali slabši od referenčne ocene večinskega klasifikatorja. V obdobjih po velikih izgubah je bilo opaziti zmanjšanje pozitivnega sentimenta. Stopnja negativnega sentimenta je bila vedno manjša od pozitivnega sentimenta. Stopnja negativnega sentimenta je bila vedno nizka, spreminjala se je samo stopnja pozitivnega sentimenta.

Izdelali smo uspešen model za trgovanje na podlagi tehnične analize in raziskali področje semantične analize. Na obeh področjih obstajajo nadaljnje izboljšave, ki niso zajete v tem delu.

Na področju tehnične analize bi lahko v model vključili več kazalnikov. V trgovalni strategiji bi lahko optimizirali upravljanje z denarjem. Namesto kupovanja največjega možnega števila delnic, bi lahko kupovali glede na podporo kazalnikov. Pri večji podpori za rast bi kupili več delnic, pri manjši podpori pa manj in tako zmanjšali tveganje in izgube ob slabih nakupih. S

podatki, ki so bili na voljo, smo se osredotočili smo se na dolgoročno trgovanje. Če bi imeli dostop do dnevnih podatkov bi lahko model razširili tudi na visokofrekvenčno trgovanje.

Največja težava pri semantični analizi je bila pomanjkanje letnih poročil delničarjem. Večina podjetij na svojih spletnih straneh ponuja dokumente za zadnjih pet let, kar je premalo za obdobje dvajsetih let, ki smo ga uporabljali. Analiza je bila zelo preprosta, saj je temeljila na štetju pojavitev besed. Napoved bi lahko izboljšali z analizo konteksta v katerem se pojavljajo besede.

Model je pokazal dobro delovanje in potencial za razširitev. S predlaganimi izboljšavami in nadaljnjim razvojem se lahko začne uporabljati kot samostojno orodje za trgovanje z vrednostnimi papirji.

# Literatura

- [1] John Murphy, Technical analysis of the financial markets : a comprehensive guide to trading methods and applications, New York Institute of Finance, New York, 1999.
- [2] Charled Kirkpatrick, Technical analysis : the complete resource for financial market technicians, FT Press, New Jersey, 2011
- [3] Simple Moving Average, <http://www.investopedia.com/terms/s/sma.asp>, Dostopano: 13.8.2017
- [4] Relative Strength Index, <http://www.investopedia.com/terms/r/rsi.asp>, Dostopano: 13.8.2017
- [5] Stochastic Oscillator, [http://stockcharts.com/school/doku.php?id=chart\\_school:technical\\_indicators:stochastic\\_oscillator\\_fast\\_slow\\_and\\_full](http://stockcharts.com/school/doku.php?id=chart_school:technical_indicators:stochastic_oscillator_fast_slow_and_full), Dostopano: 13.8.2017
- [6] Volatility, <http://www.investopedia.com/terms/v/volatility.asp>, Dostopano: 13.8.2017
- [7] Pridobivanje zgodovinskih podatkov iz Google Finance, <http://www.learn datasci.com/python-finance-part-yahoo-finance-api-pandas-matplotlib>, Dostopano: 13.8.2017
- [8] Opinion Mining, Sentiment Analysis, and Opinion Spam Detection, <https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>, Dostopano: 13.8.2017